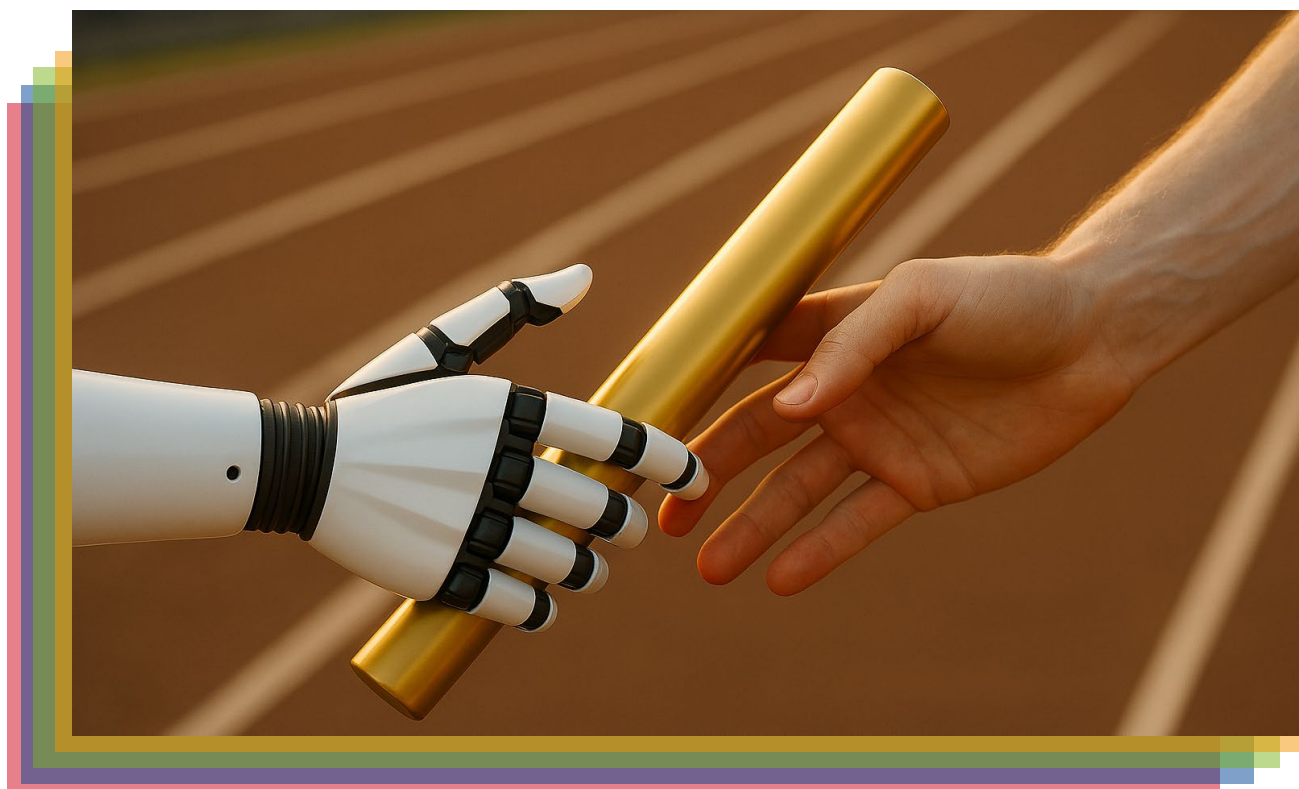


SYNTHESIS

EXPLORING GREY LITERATURE WITHIN THE NEED INITIATIVE: WHAT POTENTIAL DOES AI HAVE?



CÉDRIC SCHOONEN, LAURENCE PIRENNE, ANOUK BERGEZ, CAROLINE OBYN, IRINA CLEEMPUT,
CHARLINE MAERTENS DE NOORDHOUT, VINCENT BOUCHER



■ PREFACE

Exploring grey literature has been a considerable challenge for researchers worldwide for decades. By definition, grey literature is not included in traditional bibliographic databases such as Medline. It can also take very diverse forms: reports from public authorities, vision papers from interest groups, locally published scientific research, etc. Finding these sources is difficult and time-consuming, and requires having at least an idea of where to search.

However, searching for grey literature sources is also essential: when results are not published in peer-reviewed journals, this is usually not a coincidence. Research has shown that studies with negative results or written in languages other than English are less frequently published through traditional channels, which means it may be necessary to look for them in grey literature.

Grey literature is also particularly important for certain research questions, such as those addressed in NEED reports. To feed the NEED database, it may be necessary to make use of information on patient needs collected by patient associations, through national surveys, or through research in very specific patient groups – all sources that may not make it into peer-reviewed journals and will sometimes simply be made available on websites.

The internet has both facilitated and complicated the search for grey literature: the quest is potentially limitless, its results always incomplete. However, the development of artificial intelligence (AI) tools opens new perspectives to facilitate this process, ideally with limited bias and in a transparent and reproducible way. Is such a strategy truly feasible and, if so, how could it be implemented in practice? These are the questions this report seeks to answer, with the ultimate aim of increasing research efficiency and improving data quality.

Christophe JANSSENS
Deputy General Director

Ann VAN DEN BRUEL
General Director



■ TABLE OF CONTENT

■	SYNTHESIS	3
1	CONTEXT AND OBJECTIVES	3
1.1	THE NEED PROJECT	3
1.2	OBJECTIVES AND APPROACH.....	3
2	IDENTIFICATION AND EVALUATION OF POSSIBLE TOOLS	4
2.1	THE PROCESS	4
2.2	KEY FINDINGS	5
3	INTEGRATION STRATEGY	6
3.1	MULTIPLE TOOLS AND AN ORCHESTRATOR	6
3.2	AVANTAGES OF THE PROPOSED ORCHESTRATION FRAMEWORK	7
4	WHAT NEXT ?	7
■	RECOMMENDATIONS.....	8



■ SYNTHESIS

1 CONTEXT AND OBJECTIVES

1.1 The NEED initiative

Launched in 2024 in the context of the Belgian Presidency of the Council of the European Union, the NEED initiative (Needs Examination, Evaluation and Dissemination) aims **to identify, collect, and assess unmet patient and societal needs in a scientific and standardised manner using a structured framework**. The data collected are integrated into a centralised database accessible to all stakeholders in health care and health innovation (as well as to the general public) to help guide interventions toward areas where they are most needed.

The information included in the NEED database comes from **various sources**: surveys and individual interviews with patients, clinical expert opinions, other national or international databases, scientific literature indexed in specialised bibliographic databases, and grey literature (see Box 1).

More information on the NEED initiative and methodology can be found in the publications available on the dedicated website <https://health-needs.eu/index.php/en/>, by selecting “search > publications” in the menu.

Box 1– Grey literature: a valuable but often hard to access source

The concept of grey literature encompasses a wide range of documents produced by government bodies, universities, NGOs, agencies, and various other organisations, which may contain relevant or even essential data **but are not published in peer-reviewed scientific journals**. These may include reports from official organisations (such as KCE reports, for instance), policy, administrative or technical documents, conference proceedings, opinion pieces, strategic vision papers, etc.

Although valuable, these sources are much more difficult to identify and use than scientific articles because they are **scattered across numerous websites** in a **variety of formats** and are **not systematically compiled or indexed in centralised databases**.

As a result, researchers are often forced to manually search for grey literature on the internet or on a limited number of sites they already know. This task requires considerable time and effort, with the risk of still **missing important information** from less well-known or less visible websites.

1.2 Objectives and approach

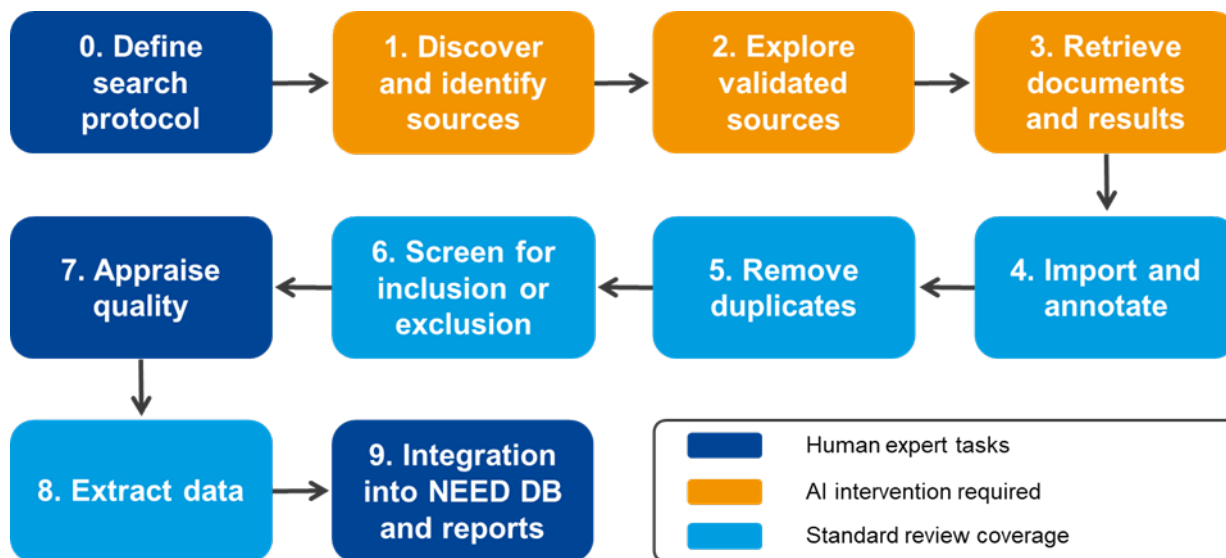
In this context, the research team investigated how artificial intelligence (AI) can contribute to identifying sources of grey literature, to the more in-depth exploration of these sources, and to the extraction of their content, with a view to accelerating the population of the NEED database. **Importantly**, the data included in the NEED database must comply with a number of legal, ethical, operational, data governance, and security requirements. These requirements are stringent and therefore not only define the scope of acceptable technical solutions but also shape the entire workflow.

As a first step, the workflow for reviewing and collecting grey literature within the NEED project was mapped (see Figure 1) in order to clarify which stages the project should focus on. Three types of tasks were retained for this study: the search for and identification of new grey literature sources, the detailed exploration of these sources in search of relevant information, and finally the extraction of documents and data.



A number of promising tools for these different tasks were then identified, categorized, compared, and assessed (first objective of the study), and a concrete approach was developed to integrate them into the workflow (second objective).

Figure 1 – Workflow for reviewing grey literature in NEED projects



The **dark blue** boxes represent tasks requiring specialized knowledge and decision-making by a human expert, the **orange** boxes tasks where the use of AI should be prioritized, and the **light blue** boxes tasks which are already covered by existing IT tools.

2 IDENTIFICATION AND EVALUATION OF POSSIBLE TOOLS

2.1 The process

To achieve the first objective of this study, the team applied a multi-step process. The first step was to identify a **wide range of tools^a** likely to provide effective support to researchers involved in the NEED initiative for obtaining data from grey literature. In total, 62 tools were identified during this phase.

The authors then classified these resources into **five categories** (see **Box 2**) based on their nature and function.

After an initial screening to eliminate those that did not meet the project's requirements, the tools were evaluated based on **22 essential or secondary criteria**. These criteria covered aspects such as functionality (ability to identify new sources or retrieve documents, types of sources the tool can access, types of data it can extract, etc.), usability (interface, technical support, configuration or integration options, etc.), and availability (availability of source code, cost, etc.), as well as specific criteria related to use in scientific projects (transparency, validation, suitability for the field of health care and grey literature, etc.).

^a In this context, "tools" refers not just to software but also to other resources, such as search engines, specialised databases, etc.

Finally, the evaluation results and the authors' expertise were used to rank the tools according to their **ability to perform the required tasks**.

AI was also used at several points in the research process to facilitate the identification of candidate tools and to enable a standardised and consistent evaluation of a large and diverse set of resources according to the needs of the NEED project, while always incorporating human verification at key decision-making stages.

Box 2 – Five categories of tools/resources

- **Databases**, registries, or “repositories” of potentially relevant content for the NEED database (articles, documents, lists of useful sources, etc.). This category is not included in the analysis, as these sources constitute potential inputs to the research workflow rather than engines to drive the process.
- AI-based search tools used **to identify websites or databases** worth investigating. Thanks to the use of large language models (LLMs), these tools can perform searches based on the meaning of a query rather than on specific keywords (e.g. “Identify reports on depression published over the past five years by government organizations in Europe”).
- Tools designed to **retrieve documents or data from a list of predefined websites**. This category includes web crawlers, which follow hyperlinks from a reference page to index content (e.g. all pages of a site), and web scrapers, which extract data from those pages. Most modern tools combine both functions.
- Tools designed to **accelerate or facilitate certain tasks** in systematic reviews (e.g. deduplication, screening, synthesis).
- Highly specialized tools that support a specific step in the workflow but do not fall into the other categories (**miscellaneous** category), such as keyword generators or query translation tools.

To meet the second objective of the study while respecting various constraints (ICT, budgetary, legal, etc.), the team developed a framework for a global solution allowing the integration of the most useful identified tools into the NEED workflow.

2.2 Key findings

Three AI search tools were selected to perform task 1 of the workflow (see **Figure 1**), i.e. the identification of potential data sources: **Exa.ai**, **Perplexity.ai**, and **Tavily**. Two web crawler/scrapper tools were selected to perform tasks 2 and 3, i.e. the exploration and retrieval of potential data: **Stormcrawler** and **Crawl4AI**.

This part of the study led to two major conclusions.

The first is that the existing offer shows clear **functional segmentation**. In other words, available tools generally target a very specific aspect of the grey literature review process (see **Figure 1**): some can identify new relevant sources, others specialise in exploring given sources or extracting data or documents, while others can accelerate later stages of data review. **No single tool is capable of covering the entire process of obtaining data for the NEED database from grey literature**. This is due to the extreme complexity of the process, driven by factors such as fragmentation and format diversity, and to the much more heterogeneous nature of metadata^b in grey literature compared to scientific literature.

^b Hidden data that provides information relating to other data, such as the author, the date of creation or the size of a document, keywords, reproduction or usage rights, etc.



The second conclusion is that the feasibility of automating a task is mainly determined by **practical constraints**. Can the tool **interact** smoothly with a broader system such as a coordination interface? Is it actively **maintained** to limit operational (e.g. technical failures) and security risks? Is it likely to encounter obstacles in accessing the data source, such as CAPTCHA protections or restrictions in terms of use? These three factors are the primary determinants of whether a tool is a good candidate.

3 INTEGRATION STRATEGY

3.1 Multiple tools and an orchestration layer

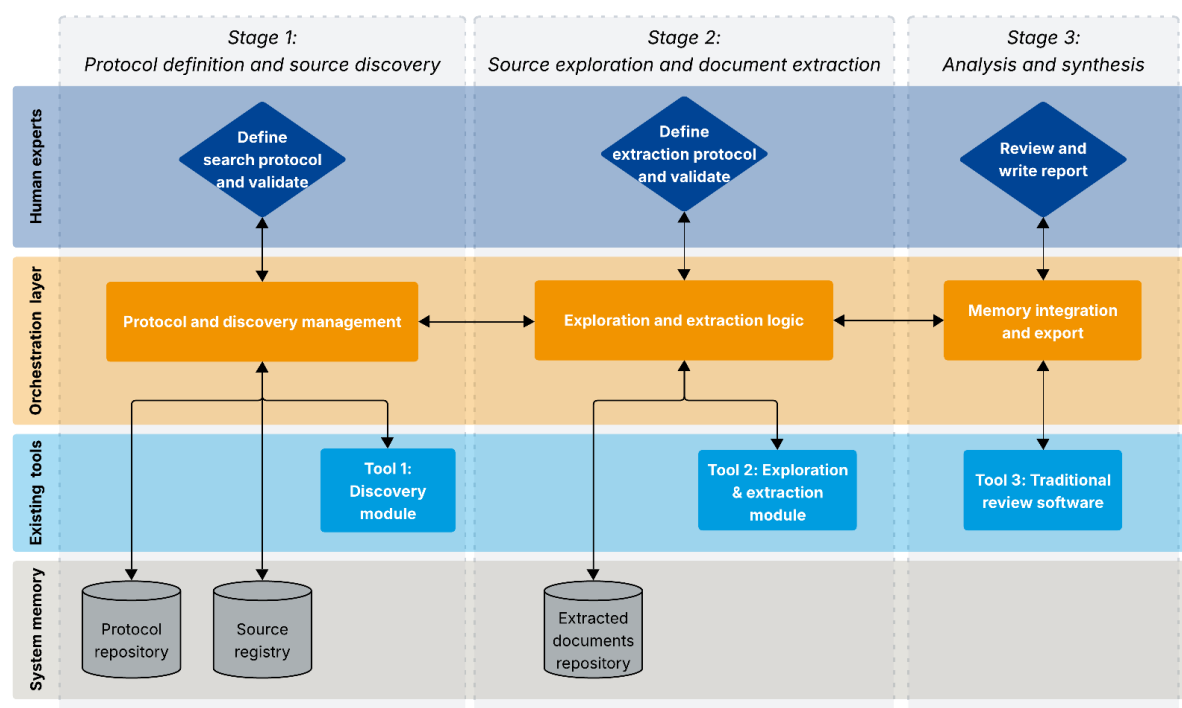
As discussed above, there is no off-the-shelf solution that fully meets the needs of the NEED initiative. The option of developing a custom-built tool was also dismissed, as this would be both extremely costly and time-consuming, delaying the availability of a solution. However, the identification and evaluation process (see above) made it possible to select several tools that could be **safely and effectively combined within a broader architecture to create a suitable solution**.

The research team therefore proposes the implementation of an “**orchestration framework**” consisting of a custom orchestration layer that leverages specialised external tools such as Exa.ai, Perplexity.ai, Tavily, Stormcrawler, and Crawl4AI. These tools are used as interchangeable modules to manage three phases of the process which cover the main tasks of grey literature research:

- protocol development and source identification,
- exploration and extraction of documents and information,
- transfer of information to the conventional research environment (human experts and standard review tools).

This system would make it possible to **automate key steps of the grey literature review process**, particularly the identification of new sources and the extraction of relevant documents, data, or information, thereby relieving human researchers of some of the most time-consuming tasks.

Figure 2 – Schematic view of the proposed orchestration framework





3.2 Advantages of the proposed orchestration framework

One of the advantages of this approach is that it ensures system **independence from external products, enhancing resilience and sustainability**, since tools can be replaced at any time depending on market developments. It would also allow KCE to retain control over protocols, validation decisions, and data provenance and recording.

Moreover, the system **standardises** the workflow and the way in which information (relating both to the collected data and to the process itself) is **documented and stored in centralised internal repositories** (see **Box 3**). This ensures **traceability** for future updates and improves efficiency by enabling the reuse of protocols, source registries, and extraction configurations. These aspects are particularly important in the context of NEED, which emphasizes transparency, traceability, and methodological consistency.

However, it is essential not to rely blindly on AI: all data obtained must be **validated by a human expert** before being added to the NEED database. The purpose of using AI tools, particularly LLMs, is to accelerate and facilitate a process that would be extremely time-consuming if carried out manually.

Box 3 – A three-part memory system

The “memory” of the proposed system is based on three separate repositories that store all data that may still be needed in the future. This includes, in particular, enabling the easy retrieval of content or sources (e.g. for control or updating purposes, or because they may be useful for another project), allowing the reuse of previously completed methodological work, and enabling the system to spontaneously reuse this information and thus improve its performance over time:

- A repository for protocol elements, which stores, for example, the methodology as well as the keywords and queries used for searches.
- A source registry that lists all identified websites, along with information on their credibility, thematic scope, etc.
- A repository for documents and other content extracted from websites, including their metadata (where possible).

4 WHAT NEXT ?

At present, the orchestration platform developed in this study remains purely theoretical. The authors recommend a **stepwise implementation**, in which the system would first be developed in a basic form and then progressively refined. This would make it possible to **quickly address the most urgent needs** (i.e. responding to data needs related to grey literature) while clarifying the technical requirements of the future platform. The system could then be gradually developed to increase its added value for research and to strengthen its technical and operational reliability, while respecting specific public sector-constraints.

In practical terms, the first step will be **to test the most promising external tools** in practice (i.e. to ensure that they are capable of performing the required tasks in the context of the NEED initiative and to assess their reliability and compatibility with institutional constraints), for example by testing them on previous NEED projects, and to compare them with one another (benchmarking). The objective is not to select a single tool, but rather to identify the most relevant combinations within a coherent architecture. Subsequently, **basic orchestration using scripts** (i.e. sequences of instructions) could be introduced to automate repetitive transfers between the different modules of the tool framework and to generate structured outputs. This would already relieve users of some time-consuming tasks, while also laying the technical foundations for subsequent stages (e.g. defining input and output data, error-handling, logging requirements). At this stage, the information to be retained would be stored in



simple shared files accessible to all researchers who may need them within KCE, or in the reference management systems already in use.

The next step would be to implement a **fully integrated platform** as described in the previous sections, offering an operational workflow for the main stages of the process (identification, validation, and exploration of sources, followed by extraction and transfer of content to conventional literature review software), as well as a centralized and controlled storage system as described in Box 3, all within a secure institutional environment.

Additional improvements or features can always be added at a later stage. However, because it structurally integrates the reuse of sources, protocols, and previous workflows stored in its three memory components, the platform is already designed to become more efficient over time.

■ RECOMMENDATIONS^c

To KCE

1. **Adopt a phased implementation model for AI-supported grey literature reviews.** AI should be introduced gradually, prioritising tools that are as unbiased as possible, so that its usefulness can be assessed at each stage. If it proves effective, AI could then be introduced through the progressive automation of specific tasks. This approach reduces the potential risks of automation and allows the methodology and governance to be adjusted over time.
2. **Position AI as a support tool for research integrated into public evidence systems, rather than as a standalone tool for exploring grey literature.** AI should assist researchers, not replace their methodological judgment or decision-making.
3. **Use AI within clear and transparent workflows.** AI can be highly useful in the literature review process, for example in identifying relevant grey literature sources or extracting usable information from diverse formats. However, it should be integrated into transparent, documented workflows subject to expert validation. Scientific accuracy and credibility depend on clear protocols, documented sources, and human oversight.
4. **Treat research protocols, validated source registries, and extracted grey literature as long-term institutional assets.** These should be stored and reused to avoid duplication of work, improve methodological consistency across studies, and make the production of future evidence more efficient and reliable.
5. **Continue to publish advances in the use of AI and make tools available to other public institutions to avoid duplication.**

^c KCE remains solely responsible for the recommendations.



COLOPHON

Title:	Exploring grey literature within the NEED initiative: what potential does AI have? - Synthesis
Authors:	Cédric Schoonen (B12 Consulting), Laurence Pirenne (B12 Consulting), Anouk Bergez (B12 Consulting), Caroline Obyn (KCE), Irina Cleemput (KCE), Charline Maertens de Noordhout (KCE), Vincent Boucher (B12 Consulting)
Senior supervisor:	Irina Cleemput (KCE)
Project manager:	Sabine Corachan (KCE)
External validators:	Benoit Macq (UCLouvain), Hubert Stoop (MyData-Trust), Alexandre Templier (Quinten)
Acknowledgements:	Rani Claerman (Sciensano), Robby De Pauw (KCE), Claudia Schoenborn (KCE), Julie Paradis (KCE), Laurence Kohn (KCE), Nicolas Fairon (KCE), Patrice Chalon (KCE), Stephan Devriese (KCE), Elena Costa (KCE), Céline Pouppez (KCE), Bertrand Chenal (B12 Consulting)
Reported interests KCE members:	KCE is a federal institution funded by INAMI/RIZIV, by the Federal Public Service of Health, Food Chain Safety and Environment, and by the Federal Public Service of Social Security. KCE's mission is to advise policymakers on decisions relating to health care and health insurance on the basis of scientific and objective research. It is expected to identify and shed light on the best possible solutions, in the context of an accessible, high-quality health care system with due regard for growing demand and budgetary constraints. KCE has no interest in companies (commercial or non-commercial, i.e. hospitals and universities), associations (e.g. professional associations, unions), individuals or organisations (e.g. lobby groups) that could be positively or negatively affected (financially or in any other way) by the implementation of the recommendations. All experts involved in the writing of the report or the peer-review process completed a declaration of interest form. Information on potential conflicts of interest is published in the colophon of this report. All members of the KCE Expert Team make yearly declarations of interest and further details of these are available upon request.
Reported interests:	None declared
Layout:	Ine Verhulst
Financial support:	The NEED project is financially supported by Belspo, under grant number RT/23/NEED
AI Disclaimer:	In the development of this work, the author(s) utilised Copilot, ChatGPT and Gemini Deep Search to assist with the wording and linguistic flow and with the research process. Following its use, the content was carefully reviewed and revised by the author(s), who accept full responsibility for the final version of the manuscript.

Disclaimer:	• The external experts were consulted about a (preliminary) version of the scientific report. Their comments were discussed during meetings. They did not co-author the scientific report and did not necessarily agree with its content. • Subsequently, a (final) version was submitted to the validators. The validation of the report results from a consensus or a voting process between the validators. The validators did not co-author the scientific report and did not necessarily all three agree with its content. • Finally, this report has been approved by common assent by the Executive Board.
--------------------	---



- Only the KCE is responsible for errors or omissions that could persist. The policy recommendations are also under the full responsibility of the KCE.

Publication date:	2 April 2026
Project number:	2023-80-2_WP2
Domain:	Needs Examination, Evaluation and Dissemination (NEED)
MeSH:	Unmet needs, needs assessment, prioritisation
NLM Classification:	W84.3
Language:	English
Format:	Adobe® PDF™ (A4)
Legal depot:	D/2026/10.273/17
ISSN:	2466-6459
Copyright:	KCE reports are published under a “by/nc/nd” Creative Commons Licence http://kce.fgov.be/content/about-copyrights-for-kce-publications .



How to refer to this document?

Schoonen C, Pirenne L, Bergez A, Obyn C, Cleemput I, Maertens De Noordhout C, Boucher V. Exploring grey literature within the NEED initiative: what potential does AI have? - Synthesis. Brussels: Belgian Health Care Knowledge Centre (KCE). 2026. KCE Reports 416C.
<https://doi.org/10.57598/R416C>.

This document is available on the website of the Belgian Health Care Knowledge Centre.